Energy Efficient Routing for Statistical Inference of Markov Random Fields

Animashree Anandkumar
ECE Department
Cornell University
Ithaca, NY 14853
aa332@cornell.edu

Lang Tong
ECE Department
Cornell University
Ithaca, NY 14853
ltong@ece.cornell.edu

Ananthram Swami
Army Research Laboratory
Adelphi
MD 20783
a.swami@ieee.org

Abstract— The problem of routing of sensor observations for optimal detection of a Markov random field (MRF) at a designated fusion center is analyzed. Assuming that the correlation structure of the MRF is defined by the nearest-neighbor dependency graph, routing schemes which minimize the total energy consumption are analyzed. It is shown that the optimal routing scheme involves data fusion at intermediate nodes and requires transmissions of two types *viz.*, the raw sensor data and the aggregates of log-likelihood ratio (LLR). The raw data is transmitted among the neighbors in the dependency graph and local contributions to the LLR are computed. These local contributions are then aggregated and delivered to the fusion center. A 2-approximation routing algorithm (DFMRF) is proposed and it has a transmission multidigraph consisting of the dependency graph and the directed minimum spanning tree, with the directions toward the fusion center.

Index Terms— Routing, Detection, Markov random fields, Graph theory.

I. INTRODUCTION

The design of routing is crucial for sensor networks, since sensors have limited transmission range and are also energy constrained. With regard to the sensor networks, the notion of cooperative routing is particularly relevant, since they are designed to accomplish a common set of objectives. For example, for a detection application, the goal is to minimize the global error probability. Moreover, for the detection application, information can be progressively fused along the route; this is advantageous since energy required for local computations is typically far less than that required for the transmission of raw data [1]. In this paper, we focus on the routing scheme for detection that minimizes the total energy consumption. This optimization problem is known as the minimum energy routing [2]. Further, we assume that the schemes achieve an optimal detection performance at a designated fusion center. If we do not allow data fusion at any intermediate nodes, then the minimum energy routing is given by the union of all shortest paths to the fusion center. However, this may not be as energy efficient as the scenario in which data fusion occurs enroute to the fusion center.

The feasibility of such data fusion, however, depends on the physical spatial signal being sensed; it is especially influenced by the correlation structure of the signal. Also, in general, these spatial random signals are acausal in contrast to the temporal signals. In the literature, the two are usually distinguished by referring to the acausal signals as random fields (RF) and to the causal signals as random processes (RP). Examples of correlated sensors include temperature and humidity sensors, and magnetometric sensors tracking a moving vehicle. Acoustic data is also rich in spatial correlations, due to the presence of echoes. In this paper, we assume that the statistical dependence or the correlations of the sensor observations is represented by a Markov random field (MRF). MRF falls under the framework of graphical models and the statistical dependence of data is defined by an undirected graph, known as the dependency graph.

A. Related work and contributions

In this paper, we use the term *data fusion* to mean that some intermediate processing of data occurs at various nodes, before reaching the fusion center. *Data aggregation* is defined as the process of computing functions (*e.g.,* sum, max) satisfying the condition that multiple incoming values at a node can be processed to a single value and forwarded to the next node. We use the term *data gathering* to mean that raw data is transmitted without any processing at any of the nodes.

An overview of routing for mobile-wireless networks can be found [1], [3], [4]. A survey of in-network processing for computing certain aggregate functions is dealt in [5], [6]. Markov random fields (MRF) are also known as conditional auto-regressions (CAR), in the seminal work of Besag [7]. For detailed explanation and examples of MRF, see [8], [9].

In [10]–[20], correlated data gathering and aggregation schemes have been considered. The works in [10]–[13] along with ours, are aggregation-driven schemes, in the sense that the routing paths are designed to benefit from data correlation. On the other hand, in routing-driven algorithms [14]–[19], data is routed to the sink through the shortest-path routing

and aggregation occurs opportunistically, whenever the data streams meet.

A related algorithm employed in the MRF framework is the belief propagation (BP), also known as the sum-product algorithm [21]. In [22], a dynamic-programming approach to resource management for object tracking is proposed. However, the possibility of data fusion, enroute, is not considered. In [23], a decision-theoretic approach to inference with single bit communication is considered. However, the network topology is predefined by a directed acyclic graph. Application-specific routing for detection or the so-called Chernoff routing, with a specific link-metric has been proposed in [24]. However, this approach assumes the model of an one-dimensional Gauss-Markov random process, which is not applicable when the nodes are on a plane.

The schemes we propose in this paper apply for a single round of sensor observations, which is different from the block-coding or the flow approach with large number of samples at every sensor. Moreover, none of the previous works incorporate the MRF model, which allows us to efficiently fuse the correlated data. The data-aggregation schemes considered in the literature (*e.g.*, [6]) are not applicable here, unless the sensor data are conditionally independent. In a related work [25], we have derived the closed-form error exponent for Neyman-Pearson detection of a Gaussian MRF with nearest-neighbor dependency graph. In this paper, we show that routing with data fusion requires transmissions of two kinds *viz.*, the raw sensor data and the aggregates of log-likelihood ratio (LLR). The raw sensor data are transmitted among the neighbors in the dependency graph and local contributions to the LLR are computed. These local contributions are then aggregated and delivered to the fusion center along the directed minimum spanning tree, with directions toward the fusion center.

Our paper is organized as follows. In section I-B, we define some graph structures. In sections II, III and IV, we provide the system model, the problem formulation and the routing algorithm. Section V concludes the paper.

B. Graph structures

Notation: Vectors and matrices are written in boldface. Random variables are in capital letters, random processes and random fields in boldface capitals and sets in calligraphic font. For two sets $\mathcal{A}$ and $\mathcal{B}$, let $\mathcal{A} \backslash \mathcal{B} = \{i : i \in \mathcal{A}, i \notin \mathcal{B}\}$.

An undirected graph $\mathcal{G}$ is a tuple $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V}$ is the vertex set and $\mathcal{E} = \{(i,j), i,j \in \mathcal{V}, i \neq j\}$ is the edge set. When i and j have an edge between them, i and j are neighbors, denoted by $i \sim j$. The neighborhood function of a node i is the set of all other nodes having an edge with it,

$$\mathcal{N}_e(i) = \{j \in \mathcal{V} : j \neq i, (i,j) \in \mathcal{E}\}. \tag{1}$$

The number of neighbors of a node i is called its degree, denoted by $\mathrm{Deg}(i)$. Let r_e denote the Euclidean length of edge e. The set of nodes of degree one are known as the leaves, denoted by $\mathcal{L}_f(\mathcal{G})$; the set of nodes $\mathcal{V} \backslash \mathcal{L}_f(\mathcal{G})$ are known as the

internal nodes. Let $\mathrm{Prune}(\mathcal{G})$ denote an operation that removes all the leaves from $\mathcal{G}$ *i.e.*,

$$\mathrm{Prune}(\mathcal{G}) \overset{\Delta}{=} (\mathcal{V} \backslash \mathcal{L}_f(\mathcal{G}), \mathcal{E}). \tag{2}$$

For a directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, we denote the edges (arcs) by $\mathcal{E} = \{<i,j>, i,j \in \mathcal{V}, i \neq j\}$, where the direction of the edge $<i,j>$ is from i to j. j is said to be a direct successor of i, and i is said to be a direct predecessor of j.

The nearest-neighbor function of a node $i \in \mathcal{V}$ on Euclidean space[1] is defined as

$$\mathrm{nn}(i) \overset{\Delta}{=} \arg \min_{j \in \mathcal{V}, j \neq i} \mathrm{dist}(i,j), \tag{3}$$

where $\mathrm{dist}(\cdot, \cdot)$ is the Euclidean distance. We assume that the inter-node distances are unique, thereby ensuring that (3) is well-defined. The nearest-neighbor (undirected) graph $\mathrm{NNG}(\mathcal{V})$ is given by

$$(i,j) \in \mathrm{NNG}(\mathcal{V}) \iff i = \mathrm{nn}(j) \text{ or } j = \mathrm{nn}(i). \tag{4}$$

NNG has a number of important properties [26]. It is acyclic with a maximum node degree of five, for unique inter-node distances.

The minimum spanning tree $\mathrm{MST}(\mathcal{V})$ over a point set $\mathcal{V}$ in Euclidean space is defined as the tree of minimum total length that spans all the points in $\mathcal{V}$ [27]. Efficient algorithms for finding the MST have been proposed in [28], [29] and can be implemented at the nodes in a distributed manner. Also, the MST is invariant under any monotone transformation of edge lengths [27]. The directed minimum spanning tree $\mathrm{DMST}(\mathcal{V})$ is defined in this paper, as the MST with directions toward the fusion center.

The unit disk graph $\mathrm{UDG}(\mathcal{V})$ is defined such that there is an edge between any two nodes in $\mathcal{V}$ if their Euclidean distance is at most one.

II. SYSTEM MODEL

A. Network model

We assume the presence of a medium-access control (MAC), which eliminates collisions or interferences among the nodes. We assume that all real numbers are quantized with sufficiently high precision so that we can ignore the quantization error. We assume that all the nodes function as both sensors and routers. By proper scaling, we assume that all the nodes have the same maximum transmission range equal to one unit and have a disk as the transmission region, with perfect data reception within this disk. Thus, the unit disk graph (UDG) represents the set of possible connections between the nodes. We assume that the UDG is connected; this implies that communication is feasible, via a multi-hop route, between any two nodes in the network. We assume that power control is carried out at the nodes and the transmission

---

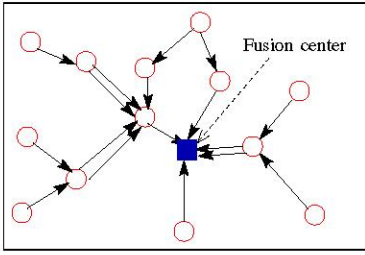[1]We assume that the nodes are on a plane.

Figure 1: Routing to the designated fusion center is represented through a multidigraph contained in unit disk graph. Each arc represents the transmission of one real number.

powers are adjusted according to their neighbors' positions. We assume that a given transmission from a node is to a single receiver (unicast mode).

B. Energy model

In order to include the energy constraint in our formulation, we assume that the energy consumed by a node i can be represented by the sum of a constant processing energy $C_p > 0$ and the transmission energy[2] [30]. It is assumed that the energy required for the transmission of any real number from i to j is the same, and is denoted by $C_{i,j}$. Further, we assume that the processing energy C_p at every node is constant and independent of the number of transmissions from the node. The path loss in transmission normally depends on the heights of the transmit antennas, as well as the transmitter-receiver separation. In this paper, we assume that all the nodes have similar antenna heights so that we can focus on the path loss that is distance-dependent. Therefore, $C_{i,j}$ is given by

$$C_{i,j} = C_t |\text{dist}(i,j)|^\nu, 2 \leq \nu \leq 4, \tag{5}$$

where C_t is a constant, $\text{dist}(i,j)$ is the inter-node distance and ν is the channel path-loss exponent.

The set of transmissions from a set of nodes $\mathcal{V}$ can be represented as a multidigraph $\mathcal{G}_t = (\mathcal{V}, \mathcal{E}_t)$, defined as a directed graph with multiple arcs and no loops. See Fig.1. Let $\mathrm{C}(\mathcal{G}_t)$ denote the total transmission energy given by

$$\mathrm{C}(\mathcal{G}_t) \triangleq C_t \sum_{e \in \mathcal{E}_t} r_e^\nu, \tag{6}$$

where r_e is the Euclidean edge length. We further require that $\mathcal{G}_t \subset \mathrm{UDG}(\mathcal{V})$ (when multiple arcs are counted as a single edge and the directions are ignored), for the transmissions to be feasible.

C. Data model : Markov random fields

We assume that the sensors collect data from a Markov random field (MRF). A MRF satisfies special conditional

independence properties based on an underlying graph structure, known as the dependency graph[3]. A simple example is the first-order auto-regressive process, where the notion of conditional independence is based on an ordering of the variables. Here, variables can be represented as a linear graph such that given a variable, the two parts of the graph, obtained by removing the variable, are conditionally independent. However, a spatial random field has a richer set of conditional independencies, requiring a more general definition.

Definition 1 (MRF): Let $\mathcal{V} = \{1, \ldots, n\}$ be a set of n nodes on the plane. The random vector of observation samples $\mathbf{Y}_n \triangleq [Y_1, \ldots, Y_n]^T$ is a Markov random field with an (undirected) dependency graph $\mathcal{G}_d = (\mathcal{V}, \mathcal{E}_d)$ if for a continuous density f, $f(\mathbf{Y}_n) > 0$ (positivity condition) and $\forall i \in \mathcal{V}$,

$$Y_i \perp \mathbf{Y}_{\mathcal{V} \backslash \{i, \mathcal{N}_e(i)\}} | \mathbf{Y}_{\mathcal{N}_e(i)}, \tag{7}$$

where $\perp$ denotes conditional independence.

We now state the famous Hammersley-Clifford theorem [31], which gives a relationship between the dependency graph and the joint density of a MRF, by proving the equivalence of the MRF and the Gibbs random field.

Theorem 1 (Hammersley-Clifford theorem): For a MRF $\mathbf{Y}_n$ with dependency graph $\mathcal{G}_d = (\mathcal{V}, \mathcal{E}_d)$, the joint PDF f can be expressed as

$$\log f(\mathbf{Y}_n; \mathcal{G}_d) = Z + \sum_{c \in \mathcal{C}} \Psi_c(\mathbf{Y}_c), \ Z \triangleq e^{-\int_{\mathbf{Y}_n} \prod_{c \in \mathcal{C}} \Psi_c(\mathbf{Y}_c)}, \tag{8}$$

where $\mathcal{C}$ is the set of all cliques in $\mathcal{G}_d$ and Ψ_c are known as the clique potentials.

Thus, the collection of conditional-independence relations in (7) results in the factorization of the joint PDF into a product of functions, each of which depends on a small set of variables. This significantly reduces the complexity of the problem.

We now make the simplification that there are only pairwise interactions or 2-clique potentials (Besag's auto-model [7]). The expression now simplifies to

$$\log f(\mathbf{Y}_n; \mathcal{G}_d) = Z + \sum_{(i,j) \in \mathcal{E}_d} \Psi_{i,j}(Y_i, Y_j) + \sum_{i \in \mathcal{V}} \Psi_i(Y_i). \tag{9}$$

We assume that local information is available: every node knows its edge and node potentials of the MRF, and its edge-lengths.

III. PROBLEM FORMULATION

We now specify the type of inference problems that can be addressed in our framework. We consider the detection of binary hypothesis (null hypothesis $\mathcal{H}_0$ and alternative hypothesis $\mathcal{H}_1$) of a MRF $\mathbf{Y}_n$ under the following variations:

1) *Detection of dependency graph:* $\mathcal{H}_0$: $\mathcal{G}_{d,0} = (\mathcal{V}, \mathcal{E}_{d,0})$ vs. $\mathcal{H}_1$: $\mathcal{G}_{d,1} = (\mathcal{V}, \mathcal{E}_{d,1})$,

2) *Detection of parameter* θ: $\mathcal{H}_0 : \mathcal{G}_d, \theta = \theta_0$ vs. $\mathcal{H}_1 : \mathcal{G}_d, \theta = \theta_1$, where θ is independent of $\mathcal{G}_d$.

For the first category of problems, we define $\mathcal{G}_d = (\mathcal{V}, \mathcal{E}_d)$, with $\mathcal{E}_d \triangleq \mathcal{E}_{d,0} \cup \mathcal{E}_{d,1}$.

Let $f(\mathbf{Y}_n | \mathcal{V}; \mathcal{H}_j)$ be the conditional PDF of the observations given the point set $\mathcal{V}$ under hypothesis j. The optimal decision-rule under both NP and Bayesian formulations is a threshold test based on the log-likelihood ratio (LLR). For hypothesis-testing problems that fall into the above-mentioned categories, from (9), the LLR can be written as

$$
\begin{aligned}
\mathrm{LLR}(\mathbf{Y}_n, \mathcal{G}_d) &\triangleq \log \frac{f(\mathbf{Y}_n | \mathcal{V}; \mathcal{H}_0)}{f(\mathbf{Y}_n | \mathcal{V}; \mathcal{H}_1)}, &\qquad(10)\\
&= Z' + \sum_{(i,j) \in \mathcal{E}_d} \Phi_{i,j}(Y_i, Y_j) \\
&\quad + \sum_{i \in \mathcal{V}} \Phi_i(Y_i), &\qquad(11)
\end{aligned}
$$

where Z' is a constant and the functions Φ are given by the difference of potentials under the two hypotheses.

A. Nearest-neighbor dependency graph

A common approach to specifying a MRF is through a neighborhood rule for the dependency graph and then to specify the potential function between these neighbors. We assume that the dependency graph $\mathcal{G}_d$ is a *proximity graph*, where the edges are included, based on the local point configuration, according to some specified rule [32]. The simplest proximity graph is the nearest-neighbor graph (NNG). The nearest-neighbor relation has been used in the literature in several areas of applied science, including the social sciences, geography and ecology, where proximity data is important (*e.g.*, [33]). In this paper, we assume that the dependency graph $\mathcal{G}_d$ in (11) is the nearest-neighbor graph.

B. Minimum energy routing

The aim of optimal routing for detection is to ensure the delivery of the LLR, given by (11), to the designated fusion center v_0, while minimizing the total energy consumption. Since we have assumed a constant processing energy C_p and that data from all the sensors is fused, this reduces to minimizing the sum of power-weighted edges of the transmission multidigraph $\mathcal{G}_t$, given by $\mathrm{C}(\mathcal{G}_t)$ in (6). As mentioned in the previous section, we further require that $\mathcal{G}_t \subset \mathrm{UDG}(\mathcal{V})$ (when multiple arcs are counted as a single edge and the directions ignored), for the transmissions to be feasible. Let $l(i; \mathcal{G}_t)$ be the value at node i at the end of a specified sequence of transmissions that are in $\mathcal{G}_t$.

Our goal is to find the optimal transmission multidigraph $\mathcal{G}^*$ such that

$$
\mathcal{G}^*(\mathcal{V}) = \arg \min_{\mathcal{G}_t \subset \mathrm{UDG}(\mathcal{V})} \mathrm{C}(\mathcal{G}_t(\mathcal{V})), \qquad(12)
$$

subject to the constraint that there exists a sequence of transmissions in $\mathcal{G}_t$ such that

$$
l(v_0; \mathcal{G}_t) = \mathrm{LLR}(\mathbf{Y}_n; \mathcal{G}_d), \qquad(13)
$$

where $\mathcal{G}_d$ is the dependency graph of MRF and LLR is given by (11). We stress that $\mathcal{G}_t$ is associated with the transmissions in contrast to $\mathcal{G}_d$, which is associated with the statistical dependence of data.

For the special case of conditionally-independent observations at the nodes ($\Phi_{i,j} \equiv 0$), the LLR in (11) is additive in the sensor data. This is the aggregation problem addressed in the literature (*e.g.*, [15]). In this case, the minimum energy aggregation graph is given by the directed minimum spanning tree (DMST), with the directions toward the fusion center.

C. Two kinds of transmission

For a general MRF, the LLR in (11) is not a sum function of the sensor data. In order to compute the edge potentials of the LLR, of the form $\Phi(Y_i, Y_j)$, access to both observations Y_i and Y_j are required. Therefore, there are transmissions of two types, *viz.*, the raw sensor data and the LLR aggregates. We formally define the two types of transmissions below.

Definition 2: The *data-transmission graph* DTG($\mathcal{V}$) is defined as the transmission subgraph that consists of transmissions of the raw sensor data. The *likelihood-aggregation graph* AG($\mathcal{V}$) is defined as the transmission subgraph that consists of transmissions of the aggregates of the log-likelihood ratio. The nodes that process data from other nodes *i.e.*, those in the likelihood-aggregation graph, are known as the set of *aggregators*, denoted by $\mathcal{V}_{\mathrm{AG}}$.

The two transmission subgraphs DTG($\mathcal{V}$) and AG($\mathcal{V}$) are not independent, since under the constraint given in (13), modifying the structure of one influences the other. For example, if a node has received the observations from all its neighbors of the dependency graph $\mathcal{G}_d$, then it needs to transmit only the LLR aggregate and not its own observation. Therefore, data fusion in MRF is a combination of data gathering and aggregation.

IV. DATA-FUSION ALGORITHM

In this section, we provide a data fusion scheme for inference of Markov random field with an approximation bound of two. See algorithms 1, 2, and Fig. 2. We specify the data-transmission graph DTG($\mathcal{V}$) in algorithm 1. The data-transmission graph DTG($\mathcal{V}$) needs to ensure that all the edge potentials of the LLR in (11), of the form $\Phi_{i,j}(Y_i, Y_j)$, are computed on transmitting the raw sensor data along the edges of DTG($\mathcal{V}$). Moreover, these edge potentials exist only between the neighbors in the dependency graph $\mathcal{G}_d$. Therefore, one possible DTG($\mathcal{V}$) is $\mathcal{G}_d$ itself, with arbitrarily assigned directions. But, in algorithm 1 (step 4), we specify that transmissions occur from the leaves to their internal neighbors. This results in energy savings since the leaves do not need to participate in the aggregation process. For other nodes (step 5), however, the assignment of directions are still arbitrary. The nearest-neighbor dependency graph is sparse and therefore, raw data from a node needs to be transmitted only to a
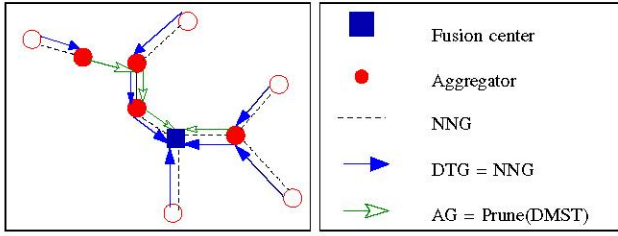
Figure 2: Schematic of DFMRF algorithm.

Algorithm 1: Data-transmission graph DTG($\mathcal{V}$)

Input: Node set $\mathcal{V}$
1. NNG($\mathcal{V}$) = Nearest-neighbor dependency graph
2. MST($\mathcal{V}$) = Minimum spanning tree
3. $\mathcal{L}_f(\mathcal{G})$ = Set of leaf nodes of graph $\mathcal{G}$
4. **if** $i \in \mathcal{L}_f(MST)$ *and* $(i,j) \in MST$ **then**
   $< i,j > \in$ DTG($\mathcal{V}$)
5. **else if** $i,j \notin \mathcal{L}_f(MST)$ *and* $(i,j) \in NNG, i < j$ **then**
   $< i,j > \in$ DTG($\mathcal{V}$)
   Output: Data-transmission graph DTG($\mathcal{V}$)

small set of neighbors (at most five). Moreover, unlike the aggregation graph AG($\mathcal{V}$), which needs to ensure that the LLR is available at the fusion center, the construction of the DTG($\mathcal{V}$) can be implemented in a localized manner.

In algorithm 2, we specify the complete data fusion mechanism. During the data-transmission phase (steps 5-6), transmissions of raw data occurs along the DTG($\mathcal{V}$). At the end of this phase, every internal node i computes and stores its local contribution $m(i)$, given by

$$m(i) = \Phi_i(Y_i) + \sum_{<j,i>\in\text{DTG}(\mathcal{V})} \Phi_{i,j}(Y_i, Y_j) + \sum_{<j,i>\in\text{DTG}(\mathcal{V}), j\notin \mathcal{V}_{\text{AG}}} \Phi_j(Y_j). \qquad (14)$$

This local contribution is then aggregated and delivered to the fusion center (steps 7-16), along the aggregation graph AG($\mathcal{V}$), given by Prune(DMST($\mathcal{V}$)) *i.e.*, the directed minimum spanning tree, with leaves removed. On receiving the aggregates from all its direct predecessors in AG($\mathcal{V}$), each node $i \neq v_0$ combines them with its local contribution and transmits $l(i)$ to its direct successor, given by

$$l(i) = \sum_{<j,i>\text{AG}(\mathcal{V})} l(j) + m(i). \qquad (15)$$

At the end of the aggregation process, the LLR is available at the fusion center and is given by $l(v_0)$.

Thus, the transmission multidigraph of the DFMRF algorithm is given by the combination of DTG($\mathcal{V}$) and AG($\mathcal{V}$),

$$\text{DFMRF}(\mathcal{V}) \triangleq \text{DTG}(\mathcal{V}) \cup \text{AG}(\mathcal{V}). \qquad (16)$$

The total transmission energy is given by

$$\text{C}(\text{DFMRF}(\mathcal{V})) = \text{C}(\text{NNG}(\mathcal{V})) + \text{C}(\text{Prune}(\text{MST}(\mathcal{V}))). \qquad (17)$$

The DFMRF algorithm requires a fixed energy and bandwidth at nodes: every node has at most six transmissions. This is because every node transmits its own data to its direct successors in DTG($\mathcal{V}$), whose number is at most five. Similarly, every node transmits its aggregate to its unique direct successor in AG($\mathcal{V}$).

Algorithm 2: Data fusion of Markov random field at a fusion center v_0 (DFMRF)

Input: DTG($\mathcal{V}$) from algorithm 1
1. DMST($\mathcal{V}$) = Minimum spanning tree, directions toward fusion center v_0
2. $\mathcal{L}_f(\mathcal{G})$ = Set of leaf nodes of graph $\mathcal{G}$
3. Prune($\mathcal{G}$) = Graph obtained by removing leaves of $\mathcal{G}$
4. AG($\mathcal{V}$) = Prune(DMST($\mathcal{V}$)) is aggregation graph
5. **foreach** $< i,j > \in$ DTG($\mathcal{V}$) **do** Transmit Y_i from i to j
6. **foreach** $i \notin \mathcal{L}_f(DTG)$ **do** Compute $m(i)$ from (14)
7. **foreach** *Node* i **do** Initialize $Tx(i) = false$
8. **foreach** *Node* $i \in \mathcal{L}_f(AG)$ **do** Set $l(i) = m(i)$
9. Transmit $l(i)$ to $k :< i,k > \in$ AG($\mathcal{V}$)
10. Set $Tx(i) = true$
11. **while** $\exists$ *Node* $i \neq v_0$ with $Tx(i) = false$ **do**
12.   **foreach** *Node* i *with* $Tx(i) = false$ *and* $Tx(j) = true, \forall j :< j,i > \in AG(\mathcal{V})$ **do** Compute $l(i)$ from equation (15)
13.   Transmit $l(i)$ to $k :< i,k > \in$ AG($\mathcal{V}$)
14.   Set $Tx(i) = true$
15. **end**
16. Compute $l(v_0)$ from equation (15)
17. **return** $LLR(\mathbf{Y}_n; \mathcal{V}) = l(v_0)$

A. Performance bounds

In this section, we quantify the performance of the DFMRF algorithm. We first provide bounds for the optimal graph $\mathcal{G}^*$.

Lemma 1 (Bounds for $\mathcal{G}^*$): The optimal transmission multidigraph $\mathcal{G}_*$ in (12) satisfies the following bounds:

$$\text{C}(\text{MST}(\mathcal{V})) \leq \text{C}(\mathcal{G}_*(\mathcal{V})) \leq \text{C}(\text{DFMRF}(\mathcal{V})). \qquad (18)$$

The lower bound is achieved iff. the sensor data are independent, conditioned on either hypotheses.

Proof : The lower bound holds since a pure aggregation scheme is not adequate to deliver the LLR to the fusion center, unless the LLR is additive in the sensor data. This happens only in the case of conditionally-independent data. The upper bound holds since DFMRF($\mathcal{V}$) satisfies all the constraints: it ensures the delivery of the LLR to the fusion center and DFMRF($\mathcal{V}$) $\subset$ UDG($\mathcal{V}$) (when multiple arcs are counted as one edge and the directions are ignored). This is due to the fact

that $NNG(\mathcal{V}) \subset MST(\mathcal{V}) \subset UDG(\mathcal{V})$, whenever $UDG(\mathcal{V})$ is connected. □

We now provide the result on the approximation bound for the DFMRF algorithm.

Theorem 2 (Approximation): DFMRF is a 2-approximation algorithm *i.e.,*

$$\frac{C(DFMRF(\mathcal{V}))}{C(\mathcal{G}_*(\mathcal{V}))} \leq 2, \tag{19}$$

over all non-empty $\mathcal{V} \subset \Re^2$.

Proof : We have $C(NNG(\mathcal{V})), C(Prune(MST(\mathcal{V}))) \leq C(MST(\mathcal{V}))$. From (17), $C(DFMRF(\mathcal{V})) \leq 2C(MST(\mathcal{V}))$. From lemma 1, $C(MST(\mathcal{V})) \leq C(\mathcal{G}_*(\mathcal{V}))$. □

V. CONCLUSION

In this paper, we considered the problem of minimum energy routing that is optimized for inference of Markov random fields. We exploited the correlation structure of the sensor data defined by the nearest-neighbor dependency graph to efficiently fuse data before delivering it to the fusion center. We proposed a two-approximation DFMRF algorithm, which is based on the minimum spanning tree, with attractive features such as bounded energy and bandwidth at every node, and single-hop transmission of raw data. Algorithms with better approximation bounds are currently under investigation.

We have made a number of simplifying assumptions. Possible extensions include considering probabilistic reception of data, relaxing the disk assumption for communications, balancing the energy drain in the network and considering other dependency graphs. We have also not exploited the broadcast nature of the wireless medium. We have also not considered the interplay between the energy and time required to fuse all the data and the quality of the resulting decision at the fusion center.

REFERENCES

[1] Akkaya K and Younis M, "A Survey of Routing Protocols in Wireless Sensor Networks," *Elsevier Ad Hoc Network Journal*, vol. 3/3, pp. 325–349, 2005.

[2] P.J. Wan, G. Călinescu, X.Y. Li, and O. Frieder, "Minimum-Energy Broadcasting in Static Ad Hoc Wireless Networks," *Wireless Networks*, vol. 8, no. 6, pp. 607–617, 2002.

[3] X. Hong, K. Xu, and M. Gerla, "Scalable routing protocols for mobile ad hoc networks," *Network, IEEE*, vol. 16, no. 4, pp. 11–21, 2002.

[4] J.N. Al-Karaki and A.E. Kamal, "Routing techniques in wireless sensor networks: a survey," *Wireless Communications, IEEE [see also IEEE Personal Communications]*, vol. 11, no. 6, pp. 6–28, 2004.

[5] A. Giridhar and P.R. Kumar, "Toward a theory of in-network computation in wireless sensor networks," *Communications Magazine, IEEE*, vol. 44, no. 4, pp. 98–107, 2006.

[6] R. Rajagopalan and P.K. Varshney, "Data aggregation techniques in sensor networks: A survey," *Communications Surveys & Tutorials, IEEE*, vol. 8, no. 4, pp. 48–63, 2006.

[7] J. Besag, "Spatial interaction and the statistical analysis of lattice systems," *Journal of the Royal Statistical Society*, vol. 36, no. B, pp. 192–225, 1974.

[8] S.L. Lauritzen, *Graphical models: Clarendon Press*, Clarendon Press, 1996.

[9] X. Guyon, *Random Fields on a Network: Modeling, Statistics, and Applications*, Springer, 1995.

[10] A. Goel and D. Estrin, "Simultaneous Optimization for Concave Costs: Single Sink Aggregation or Single Source Buy-at-Bulk," *Algorithmica*, vol. 43, no. 1, pp. 5–15, 2005.

[11] R. Cristescu, B. Beferull-Lozano, M. Vetterli, and R. Wattenhofer, "Network correlated data gathering with explicit communication: NP-completeness and algorithms," *IEEE/ACM Transactions on Networking (TON)*, vol. 14, no. 1, pp. 41–54, 2006.

[12] P. von Rickenbach and R. Wattenhofer, "Gathering correlated data in sensor networks," in *Proc. of the 2004 joint workshop on Foundations of mobile computing*, 2004, pp. 60–66.

[13] H. Luo, J. Luo, and Y. Liu, "Energy efficient routing with adaptive data fusion in sensor networks," in *Proc. of 2005 joint workshop on Foundations of mobile computing.* 2005, pp. 80–88, ACM Press New York, NY, USA.

[14] W. B. Heinzelman, A. P. Chandrakasan, and H. Balakrishnan, "An application-specific protocol architecture for wireless microsensor networks," *IEEE Trans. on Wireless Communications*, vol. 1, no. 4, pp. 660–670, Oct. 2002.

[15] B. Krishnamachari, D. Estrin, and S. Wicker, "Modeling data centric routing in wireless sensor networks," in *Proc. of IEEE INFOCOM*, New York, NY, June 2002.

[16] A. Scaglione and S. Servetto, "On the Interdependence of Routing and Data Compression in Multi-Hop Sensor Networks," in *Proc. MobiCom 2002*, Atlanta, Georgia, September 2002.

[17] W. Zhang and G. Cao, "Optimizing tree reconfiguration for mobile target tracking in sensor networks," in *INFOCOM 2004. Twenty-third Annual Joint Conf.*, 2004, vol. 4.

[18] C. Intanagonwiwat, D. Estrin, R. Govindan, and J. Heidemann, "Impact of network density on data aggregation in wireless sensor networks," in *Distributed Computing Systems, 2002. Proceedings. 22nd International Conference on*, 2002, pp. 457–458.

[19] C. Intanagonwiwat, R. Govindan, and D. Esterin, "Directed Diffusion : A Scalable and Robust Paradigm for Sensor Networks," in *Proc. 6th ACM/Mobicom Conference*, Boston,MA, 2000, pp. pp 56–67.

[20] J.P. Chou and D.K. Ramachandran, "A distributed and adaptive signal processing approach to reducing energy consumption in sensor networks," in *INFOCOM, IEEE*, 2003.

[21] J. Pearl, *Probabilistic Resoning in Intelligent Systems—Networks of Plausible Inference*, Morgan Kaufmann, 1988.

[22] JL Williams, JW Fisher III, and AS Willsky, "An Approximate Dynamic Programming Approach to a Communication Constrained Sensor Management Problem," in *Information Fusion, 2005 7th International Conference on*, 2005, vol. 1.

[23] O.P. Kreidl and A.S. Willsky, "Inference with Minimal Communication: a Decision-Theoretic Variational Approach," in *Advances in Neural Information Processing Systems*, 2006.

[24] Y. Sung, S. Misra, L. Tong, and A. Emphremides, "Cooperative Routing for Signal Detection in Large Sensor Networks," *submitted to IEEE J. Select. Area Comm.*, Jan. 2006.

[25] A. Anandkumar, L. Tong, and A. Swami, "Detection of Gauss-Markov random fields with nearest-neighbor dependency," in *Proc. of ICASSP'07*, Hawaii, USA, April 2007.

[26] D. Eppstein, M. S. Paterson, and F. F. Yao, "On Nearest-Neighbor Graphs," *Discrete and Computational Geometry*, vol. 17, pp. 263–282, 1997.

[27] R.S. Preparata, F.P. Preparata, and M.I. Shamos, *Computational Geometry: An Introduction*, Springer, 1993.

[28] P. Humblet, "A Distributed Algorithm for Minimum Weight Directed Spanning Trees," *IEEE Trans. on Communications*, vol. 31, no. 6, pp. 756–762, 1983.

[29] RG Gallager, PA Humblet, and PM Spira, "A Distributed Algorithm for Minimum-Weight Spanning Trees," *ACM Tran on Prog. Languages and Sys. (TOPLAS)*, vol. 5, no. 1, pp. 66–77, 1983.

[30] A. Ephremides, "Energy concerns in wireless networks," *IEEE Wireless Communications*, , no. 4, pp. 48–59, August 2002.

[31] J.M. Hammersley and P. Clifford, "Markov fields on finite graphs and lattices," *Unpublished manuscript*, 1971.

[32] L. Devroye, "The expected size of some graphs in computational geometry," *Computers & mathematics with applications*, vol. 15, no. 1, pp. 53–64, 1988.

[33] R.K. Pace and D. Zou, "Closed-Form Maximum Likelihood Estimates of Nearest Neighbor Spatial Dependence," *Geographical Analysis*, vol. 32, no. 2, pp. 154–172, 2000.